MACHINE LEARNING FOR PERFORMANCE IMPROVEMENT OF PERIODIC NFT-BASED COMMUNICATION SYSTEM

Oleksandr Kotlyar, Morteza Kamalian-Kopae, Jaroslaw E. Prilepsky, Maryna Pankratova, Sergei K. Turitsyn*

Aston Institute of Photonic Technologies, Aston University, Birmingham, UK
**E-mail: o.kotlyar1@aston.ac.uk*

Keywords: Machine Learning based DSP for optical transmission, Transmission system modelling, Schemes for impairment mitigation increasing data throughout and/or mutual information

Abstract

We compare performance of several machine learning methods, including support vector machine, k-nearest neighbours, k-means clustering, and Gaussian mixture model, used for increasing transmission reach in the optical communication system based on the periodic nonlinear Fourier transform signal processing.

1 Introduction

There are several major factors limiting the performance of contemporary optical communication systems, including fibre attenuation, chromatic dispersion, amplifier-induced noise, and nonlinearity of fibre. In particular, the fibre nonlinearity is often considered to be the most challenging transmission impairment in the optical fibre channel [1]. Among the many methods for the nonlinearity mitigation, the nonlinear Fourier transform (NFT) has attracted some attention recently [2]. While many other approaches are aimed at diminishing the effects of nonlinearity in fibre, the NFT-based methods use it as an element of the data modulation scheme. The NFT effectively linearises the evolution of signal even at high powers within the idealised path-average model - the lossless and noiseless nonlinear Schrödinger equation (NSL). The quasi-linear evolution of nonlinear modes allows us to utilise the efficient techniques developed for linear communication in order to improve system's performance [3]. However, in the real world applications, the fibre deviates from the idealised model, i.e. the NLS with the account of losses and noise is not exactly solvable by the means of NFT, meaning that we effectively have the channel and processing mismatch bringing about system's performance degradation. However, the NFT-based transmission systems still excel the systems utilising just dispersion compensation [4]. The important unavoidable sources of NFT transmission degradation are the practical system's gain-loss profile and optical noise. While the first effect can be partially compensated for by the means of a path-averaged approach [5], the effects of latter are more involved, and the optimal detection strategy tailored to the NFT system features is yet to be designed. This work investigates the efficacy of machine learning (ML) methods in dealing with the problem of noise that degrade the performance of NFT-based systems.

We note that in recent years ML techniques have widely been used in optical communications as the effective tool for improving systems' performance [6–8]. In this work we compare potential of several ML methods to improve the performance of a periodic NFT-based optical transmission system. Supervised, k-nearest neighbours (k-NN), Gaussian mixture model (GMM), and support vector machine (SVM), and unsupervised (k-means clustering) ML methods are used to process received data. Then we compare the methods performance in terms of the achieved BER improvement for our transmission system.

In our approach we append a periodic NFT-based communication system [9, 10] with a ML block at the receiver (Rx) after the demodulation.

The path-average propagation of the slow-varying envelope $q(z,t)$ of the electromagnetic field along the optical fibre can be approximated by the NLS: $iq_z + q_{tt} + 2q|q|^2 = n(z,t)$, where z is the distance along the fibre, t represents the retarded time and n is the ASE noise. The main idea of the NFT is that the signal can be represented in the nonlinear Fourier domain through its nonlinear spectrum (NS), where the nonlinear spectral components evolve linearly. Within the NFT-based system we modulate data using the parameters of the NS modes. The inverse NFT stage is used to produce the time-domain waveform that is then launched into the fibre. At the Rx side we compensate the phase shift in the NS that occurs during propagation and retrieve our data back [11]. Although NFT is conventionally used for localised signals, it is shown that the NFT communication systems based on the periodic signal extension can be advantageous in terms of our having a smaller processing window at Rx, lower computational complexity, and the possibility of control over the signal properties [10, 12]. For the sake of brevity, details of the considered periodic NFT system is not repeated here but can be found in [9]: in particular, the NS the signals that we use for data modulation is shown in Fig. 1 of [9]. In [9] it was also reported that the effective noise in the nonlinear Fourier domain has a nonlinear dependence to the transmitted QAM symbol structure, which, in fact, is our

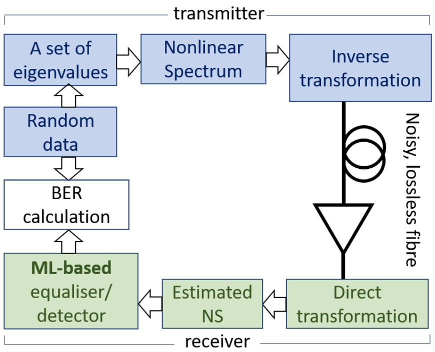

Fig. 1 The communication system based on periodic NFT and ML-based equaliser/symbol detector.

main motivation to implement ML techniques as the equalisation/detection stage. In our communication scheme, random data is mapped on a set of complex-valued eigenvalues which are used to construct a signal through the inverse transformation described in [9]. The constructed signal is then sent to the fibre undergoing linear and nonlinear distortions and perturbed by ASE noise. We assume an ideal Raman amplification providing a constant signal power over the transmission line. The distorted signal is received at the receiver and the main spectrum (comprising the eigenvalues) is calculated using through direct transformation. The received eigenvalues, distorted by the nonlinear noise-signal interaction, are fed to the ML-based equaliser/detector where an estimation of the transmitted symbol is made. The BER is then calculated by directly counting the mismatches between the sent and received QAM symbols, see Fig. 1.

We would like to mention that k-NN [13], GMM [14], SVM [15–17] and k-means clustering [17, 20] algorithms have already been successfully applied for the performance improvement of optical systems. However, to the best of our knowledge, the comparison of all these ML-methods used as the equalisers/detectors at the Rx side of optical communications systems has not been reported yet.

2 Methods and simulations results

The two simplest ML algorithms considered in our study are k-NN [21–23] and k-means clustering [21, 22]. We can use the k-NN to identify different classes of data, i.e. the particular participants (symbols) of the constellation. One of the advantages of k-NN is that this algorithm does not require any computational resources at all for the training process. Within this method, for any new unlabelled sample from testing data, we compare that new sample with each sample in the existing training set. All the distances from the testing sample to each training sample are computed. After that we determine the k nearest training data points (the nearest neighbours) and check their labels. Than, we take a majority vote from the k nearest neighbours, and the majority is the new class (the constellation symbol) we assign to the data we were asked to classify.

The unsupervised k-means clustering, or Lloyd's algorithm [24], groups the similar samples together into clusters, where each cluster defined to which constellation symbol our sample belongs. It operates with unlabelled data, i.e., the data that does not initially belong to any group (any particular constellation symbol). The algorithm partitions data into k clusters, and returns the index of the cluster to which it has assigned each sample. That index identifies the particular constellation point obtained. We use transmitted symbols for cluster centroid initialisation and also for cluster indexing. Such a initialisation provides a higher accuracy of classifying compared to *k-means++* algorithm [25] for cluster centre initialisation.

The SVM is a powerful ML method that can be used in optical communication problems for both classification and regression tasks [15, 16]. The SVM determines the support vectors and maximises their margins that are defined to be the smallest distance between the decision boundary and any of the samples. In SVM, the decision boundary is chosen to be the one for which the margin is maximised.

To classify received symbols of 64-QAM constellation, the GMM composed from the 64 linear multivariate normal density components have been employed in our work. The algorithm allocates each received point to a cluster by maximising the probability that a data point belongs to this cluster. Overall, the GMM tends to find a set of Gaussian probability distributions that describes our data points in the most accurate way.

The results of our ML methods comparison are represented via the BER vs. distance (and vs. launch power) curves to show the improvement obtained by the utilisation of supervised (k-NN, GMM, and SVM) and unsupervised (k-means clustering) ML techniques. To calculate each point in the BER figure 2^{15} symbols are used. The average signal bandwidth is 5.3 GHz and the data-rate for a 64-QAM periodic NFT-based system is around 7.2 Gb/s. Each signal is cyclically extended to the duration of the channel memory to avoid inter-symbol interference caused by chromatic dispersion.

The BER resulting from decoding by using the hard decision based on rigid rectangular symbol boundaries and the BER levels for four ML methods described above are presented in Fig. 2 for different transmission lengths ranging from 770 to 1078 km. One can see that in the range 770-924 km all studied ML methods provide a significant performance improvement. Notably, for all distances from that range, the relatively simple k-NN method operating with training samples only, which does not account for the parameters obtained from training process, gives us approximately the same BER improvement as that of the more advanced SVM algorithm. Using the proposed SVM demodulator, we achieved the BER value lower than the HD-FEC limit 3.8×10^{-3} at 1000 km, while k-NN and GMM's BER were slightly above HD-FEC. We also mention that the BER achieved by employing the other ML methods was lower than the SD-FEC threshold 2×10^{-2} [26] with k-means demonstrating the lowest performance gain but still below that value at 1000 km, while the "conventional" HD detection is above HD-FEC at approximately 950 km and above SD-FEC at 1000 km and further.

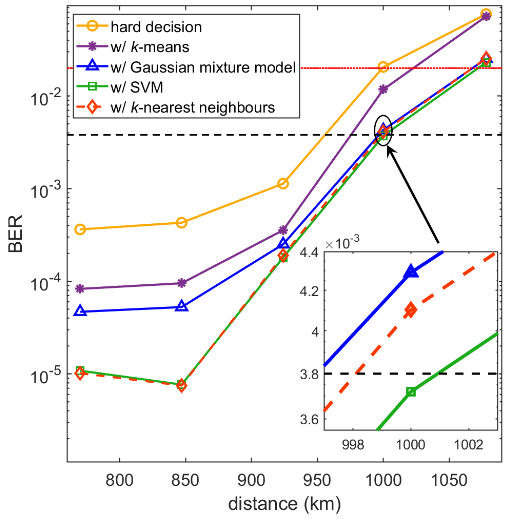

Fig. 2 The BER dependence on the transmission distance for the (almost) optimal power -7 dBm for 64-QAM periodic NFT obtained by using the conventional hard decision, k-means, GMM, SVM and k-NN based detectors (amber (circle), violet (asterisk), blue (triangle), green (square) and red (diamond) curves respectively). Horizontal black dashed and red dotted lines represents HD-FEC and SD-FEC thresholds respectively.

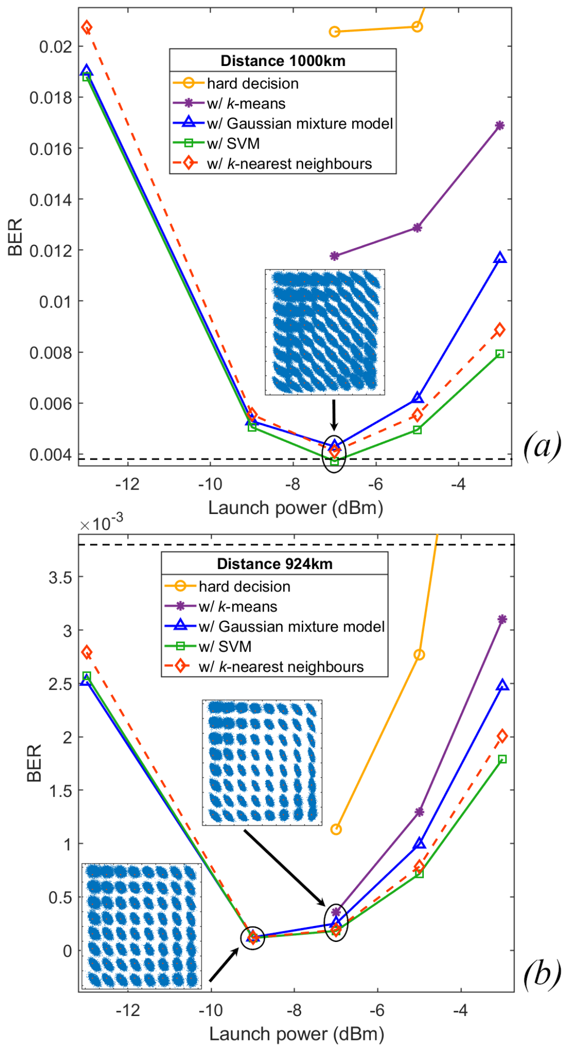

Fig. 3 BER as a function of optical launch power. Curves denote in the same way as in Fig 2 (a) at 1000 km (b) at 924 km.

Fig. 3 shows the BER curves for the received signal for 924 km and 1000 km transmission lengths as a function of the launch power. The examples of constellation diagrams obtained by system's simulation for these transmission lengths and most interesting power values are given in the insets on Fig. 3. The values of BER for the powers lower than -7 dBm achieved by conventional HD and k-means decoders much higher than the HD-FEC threshold, and so we do not show them there. As before, the best performance was rendered by the SVM for almost all power range studies, excluding the extra low power values at 924 km, and at the optimal power point the performance shown by the k-NN and GMM is very close to that of the SVM. Poor performance of k-means method can be explained by the fact that only distances to cluster centroids are taken into account for new symbols classification. Therefore, the method is failing to label correctly the symbols belonging to noncircular and located close to each other clusters. In the case of such clouds k-NN and SVM show better performance because k-NN use $8-15$ nearest to new symbol neighbours for classification and SVM creates nonlinear boundaries representing the complicated pattern of received symbols. In our opinion k-NN and SVM are more useful for compensation of nonlinearity-induced impairments while k-means and GMM can be applied to diminish the impact of ASE noise.

3 Conclusion

In this paper we compared detectors, based on three supervised (SVM, k-NN, GMM) and one unsupervised (k-means clustering) ML methods for the periodic NFT-based optical communication system. We have shown that for whole distances range addressed all studied ML methods provide a significant improvement of the system performance. At the same time, at 1000 km distance only the SVM can reduce the performance penalty below the HD-FEC threshold, though typically the simpler k-NN method is just a bit worse than more resource-demanding SVM. The attained system's performance improvements can be understood as an expansion of usable transmission distance.

4 Acknowledgements

This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie Grant Agreements No.751561 (M.P.) and No.713694 (O.K.), EPSRC project TRANSNET (EP/R035342/1) (O.K., M.K.K & S.K.T.) and the Leverhulme Trust project (RPG-2018-063) (J.E.P. & S.K.T.).

5 References

[1] Bayvel, P., Maher, R., Xu, T., et al.: 'Maximizing the optical network capacity', Phil. Trans. R. Soc. A, 2016, **374**, (2062), pp. 20140440, DOI:10.1098/rsta.2014.0440

[2] Turitsyn, S. K., Prilepsky, J. E., Le, S. T., et al.: 'Nonlinear Fourier transform for optical data processing and transmission: advances and perspectives', Optica, 2017, **4**, (3), pp. 307–322, DOI:10.1364/OPTICA.4.000307

[3] Le, S. T., Prilepsky, J. E., Turitsyn, S. K.: 'Nonlinear inverse synthesis technique for optical links with lumped amplification', Opt. Express, 2014, **23**, (7), pp. 8317–8328, DOI:10.1364/OE.23.008317

[4] Le, S. T., Aref, V., Buelow, H.: 'Nonlinear signal multiplexing for communication beyond the Kerr nonlinearity limit', Nature Photonics, 2017, **11**, (9), pp. 570–576, DOI:10.1038/nphoton.2017.118

[5] Kamalian, M., Prilepsky, J. E., Le, S. T., et al.: 'On the design of NFT-based communication systems with lumped amplification', J. Lightwave Technol., 2017, **35**, (24), pp. 5464–5472, DOI:10.1109/JLT.2017.2775105

[6] Musumeci, F., Rottondi, C., Naget, A., et al.: 'An Overview on Application of Machine Learning Techniques in Optical Networks', IEEE Communications Surveys & Tutorials, 2018, DOI:10.1109/COMST.2018.2880039

[7] Karanov, B., Chagnon, M., Thouin, F., et al.: 'End-to-End Deep Learning of Optical Fiber Communications', J. Lightwave Technology, 2018, **36**, (20), pp. 4843–4855, DOI:10.1109/JLT.2018.2865109

[8] Zibar, D., Piels, M., Jones, R., et al.: 'Machine learning techniques in optical communication', J. Lightwave Technology, 2016, **34**, (6), pp. 1442–1452, DOI:10.1109/JLT.2015.2508502

[9] Kamalian, M., Vasylchenkova, A., Prilepsky J., et al.: 'Communication System Based on Periodic Nonlinear Fourier Transform with Exact Inverse Transformation', European Conference on Optical Communication (ECOC), Rome 2018, pp. 1–3

[10] Kamalian, M., Prilepsky, J. E., Le, S. T., et al.: 'Periodic nonlinear Fourier transform for fiber-optic communications, Part I: theory and numerical methods', Opt. Express, 2016, **24**, (16), pp. 18353–18369, DOI:10.1364/OE.24.018353

[11] Thrane J., Wass J., Piels M., et al.: 'Machine Learning Techniques for Optical Performance Monitoring From Directly Detected PDM-QAM Signals', J. Lightwave Technol., 2017, **35**, (4), pp. 868–875, DOI:10.1109/JLT.2016.2590989

[12] Kamalian, M., Vasylchenkova, A., Prilepsky J., et al.: 'Signal Modulation and Processing in Nonlinear Fibre Channels by Employing the Riemann-Hilbert Problem', J. Lightwave Technol., 2018, **36**, (24), pp. 5714–5727

[13] Wang, D., Zhang, M., Fu, M., et al.: 'Nonlinearity Mitigation Using a Machine Learning Detector Based on k-Nearest Neighbors," IEEE Photonics Technology Letters, 2016, **28**, (19), pp. 2102-2105, DOI:10.1109/LPT.2016.2555857

[14] Lu, F., Peng, P., Liu, S., et al.: 'Integration of Multivariate Gaussian Mixture Model for Enhanced PAM-4 Decoding Employing Basis Expansion', 2018, Optical Fiber Communications Conference and Exposition (OFC), San Diego, CA, March 2018, pp. 1–3

[15] Nguyen, T., Mhatli, S., Giacoumidis, E., et al.: 'Fiber Nonlinearity Equalizer Based on Support Vector Classification for Coherent Optical OFDM', 2016, IEEE Photonics J., **8**, (2), pp. 1–9

[16] Giacoumidis, E., Mhatli, S., Nguyen, T., et al.: 'Comparison of DSP-based nonlinear equalizers for intrachannel nonlinearity compensation in coherent optical OFDM', 2016, Opt. Lett. **41**, (11), pp. 2509–2512

[17] Kotlyar, O., Pankratova, M., Kamalian, M., et al.: 'Unsupervised and supervised machine learning for performance improvement of NFT optical transmission', 2018 IEEE British and Irish Conference on Optics and Photonics (BICOP), London, United Kingdom, December 2018, pp. 1–4, DOI: 10.1109/BICOP.2018.8658274

[18] Zhao, T., Nehorai, A., Porat, B.: 'K-means clustering-based data detection and symbol-timing recovery for burst-mode optical receiver', 2006, IEEE Transactions on Communications, **54**, (8), pp. 1492–1501

[19] Gonzalez, N. G., Zibar, D., Yu, X., et al.: 'Optical phase-modulated radio-over-fiber links with k-means algorithm for digital demodulation of 8PSK subcarrier multiplexed signals', 2010 Conference on Optical Fiber Communication (OFC/NFOEC), San Diego, CA, March 2010, DOI:10.1364/OFC.2010.OML3

[20] Zhang, J., Chen, W., Gao, M., et al.: 'Novel Low-Complexity Fully-Blind Density-Centroid – Tracking Equalizer for 64-QAM Coherent Optical Communication Systems', 2018 Optical Fiber Communications Conference and Exposition (OFC), San Diego, CA, USA, March 2018.

[21] Harrington, P.: 'Machine Learning in Action' (Manning Publications Company, 2012)

[22] Bishop, C. M.: 'Pattern Recognition and Machine Learning' (Springer-Verlag New York, 2006)

[23] Kramer, O.: 'Dimensionality Reduction with Unsupervised Nearest Neighbors' (Springer-Verlag Berlin Heidelberg, 2013)

[24] Lloyd, S. P.: 'Least Squares Quantization in PCM', IEEE Transactions on Information Theory, 1982, **28**, (2), pp. 129–137

[25] David, A., Vassilvitskii, S.: 'K-means++: The Advantages of Careful Seeding'. SODA '07: Proceedings of the Eighteenth Annual ACM-SIAM Symposium on Discrete Algorithms, New Orleans, Louisiana, USA, January 2007, pp. 1027–1035

[26] Tzimpragos, G., Kachris, C., Djordjevic, I. B., et al.: 'A Survey on FEC Codes for 100 G and Beyond Optical Networks', IEEE Communications Surveys & Tutorials, 2014, **18**, (1), pp. 209-221, DOI:10.1109/COMST.2014.2361754